



THE UNIVERSITY *of* EDINBURGH

Edinburgh Research Explorer

Climate Change Detection and Attribution: Beyond Mean Temperature Signals

Citation for published version:

Hegerl, GC, Karl, TR, Allen, M, Bindoff, NL, Gillett, N, Karoly, D, Zhang, XB & Zwiers, F 2006, 'Climate Change Detection and Attribution: Beyond Mean Temperature Signals', *Journal of Climate*, vol. 19, no. 20, pp. 5058-5077. <https://doi.org/10.1175/JCLI3900.1>

Digital Object Identifier (DOI):

[10.1175/JCLI3900.1](https://doi.org/10.1175/JCLI3900.1)

Link:

[Link to publication record in Edinburgh Research Explorer](#)

Document Version:

Publisher's PDF, also known as Version of record

Published In:

Journal of Climate

Publisher Rights Statement:

Published in the Journal of Climate by the American Meteorological Society (2006)

General rights

Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy

The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.



Climate Change Detection and Attribution: Beyond Mean Temperature Signals

GABRIELE C. HEGERL,* THOMAS R. KARL,+ MYLES ALLEN,# NATHANIEL L. BINDOFF,@
NATHAN GILLET,& DAVID KAROLY,** XUEBIN ZHANG,++ AND FRANCIS ZWIERS##

* *Division of Earth and Ocean Sciences, Nicholas School for the Environment and Earth Sciences, Duke University, Durham, North Carolina*

+ *NOAA/National Climatic Data Center, Asheville, North Carolina*

Climate Dynamics Group, Atmospheric, Oceanic and Planetary Physics, Department of Physics, University of Oxford, Oxford, United Kingdom

@ *Antarctic Climate and Ecosystems Cooperative Research Centre, and CSIRO Marine Research, University of Tasmania, Hobart, Tasmania, Australia*

& *Climatic Research Unit, School of Environmental Sciences, University of East Anglia, Norwich, United Kingdom*

** *School of Meteorology, University of Oklahoma, Norman, Oklahoma*

++ *Climate Monitoring and Data Interpretation Division, Meteorological Service of Canada, Downsview, Ontario, Canada*

Canadian Centre for Climate Modelling and Analysis, Meteorological Service of Canada, Victoria, British Columbia, Canada

(Manuscript received 20 September 2004, in final form 20 April 2005)

ABSTRACT

A significant influence of anthropogenic forcing has been detected in global- and continental-scale surface temperature, temperature of the free atmosphere, and global ocean heat uptake. This paper reviews outstanding issues in the detection of climate change and attribution to causes. The detection of changes in variables other than temperature, on regional scales and in climate extremes, is important for evaluating model simulations of changes in societally relevant scales and variables. For example, sea level pressure changes are detectable but are significantly stronger in observations than the changes simulated in climate models, raising questions about simulated changes in climate dynamics. Application of detection and attribution methods to ocean data focusing not only on heat storage but also on the penetration of the anthropogenic signal into the ocean interior, and its effect on global water masses, helps to increase confidence in simulated large-scale changes in the ocean.

To evaluate climate change signals with smaller spatial and temporal scales, improved and more densely sampled data are needed in both the atmosphere and ocean. Also, the problem of how model-simulated climate extremes can be compared to station-based observations needs to be addressed.

1. Introduction

Evidence for an anthropogenic contribution to climate trends over the twentieth century is accumulating at a rapid pace [see Mitchell et al. (2001) and International Ad Hoc Detection and Attribution Group (2005) for detailed reviews]. The greenhouse gas signal in global surface temperature can be distinguished from internal climate variability and from the response to other forcings (such as changes in solar radiation, volcanism, and anthropogenic forcings other than greenhouse gases) for global temperature changes (e.g., Santer et al. 1996; Hegerl et al. 1997; Tett et al. 1999;

Stott et al. 2001) and also for continental-scale temperature (Stott 2003; Zwiers and Zhang 2003; Karoly et al. 2003; Karoly and Braganza 2005). Evidence for anthropogenic signals is also emerging in other variables, such as sea level pressure (Gillett et al. 2003b), ocean heat content (Barnett et al. 2001; Levitus et al. 2001, 2005; Reichert et al. 2002; Wong et al. 2001), ocean salinity (Wong et al. 1999; Curry et al. 2003), and tropopause height (Santer et al. 2003b).

The goal of this paper is to discuss new directions and open questions in research toward the detection and attribution of climate change signals in key components of the climate system, and in societally relevant variables. We do not intend to provide a detailed review of present accomplishments, for which we refer the reader to International Ad Hoc Detection and Attribution Group (2005).

Detection has been defined as the process of demon-

Corresponding author address: Gabriele Hegerl, Nicholas School of the Environment and Earth Sciences, Duke University, Durham, NC 27708.
E-mail: hegerl@duke.edu

strating that an observed change is significantly different (in a statistical sense) from natural internal climate variability, by which we mean the chaotic variation of the climate system that occurs in the absence of anomalous external natural or anthropogenic forcing (Mitchell et al. 2001). *Attribution* of anthropogenic climate change is generally understood to require a demonstration that the detected change is consistent with simulated change driven by a combination of external forcings, including anthropogenic changes in the composition of the atmosphere and internal variability, and not consistent with alternative explanations of recent climate change that exclude important forcings [see Houghton et al. (2001) for a more thorough discussion]. This implies that all important forcing mechanisms, natural (e.g., changes in solar radiation and volcanism) and anthropogenic, should be considered in a full attribution study.

Detection and attribution provides therefore a rigorous test of the model-simulated transient change. In cases where the observed change is consistent with changes simulated in response to historical forcing, such as large-scale surface and ocean temperatures, these emerging anthropogenic signals enhance the credibility of climate model simulations of future climate change. In cases where a significant discrepancy is found between simulated and observed changes, this raises important questions about the accuracy of model simulations and the forcings used in the simulations. It may also emphasize a need to revisit uncertainty estimates for observed changes.

Beyond model evaluation, a further important application of detection and attribution studies is to obtain information on the uncertainty range of future climate change. Anthropogenic signals that have been estimated from the twentieth century can be used to extrapolate model signals into the twenty-first century and estimate uncertainty ranges based on observations (Stott and Kettleborough 2002; Allen et al. 2000). This is important since there is no guarantee that the spread of model output fully represents the uncertainty of future change. Techniques related to detection approaches can also be used to estimate key climate parameters, such as the equilibrium global temperature increase associated with CO₂ doubling (“climate sensitivity”) or the heat taken up by the ocean (e.g., Forest et al. 2002) to further constrain model simulations of future climate change.

Section 2 briefly reviews methodological challenges associated with new directions in detection and attribution. Section 3 lists results and challenges in large-scale surface and atmospheric variables, while section 4 focuses on the ocean, and section 5 on impact-relevant

variables. We conclude with some recommendations in section 6.

2. Methodological challenges and data requirements

Mitchell et al. (2001) and International Ad Hoc Detection and Attribution Group (2005) give an extensive overview of detection and attribution methods. One of the most widely used, and arguably the most efficient method for detection and attribution is “optimal fingerprinting.” This is generalized multivariate regression that uses a maximum likelihood method (Hasselmann 1979, 1997; Allen and Tett 1999) to estimate the amplitude of externally forced signals in observations. The regression model attempts to represent the observed record \mathbf{y} , organized as a vector in space and time, from a set of n response (signal) patterns that are concatenated in a matrix \mathbf{X} using the linear assumption $\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{u}$. Climate change signal patterns (also called fingerprints) are usually derived from model simulations [e.g., with a coupled general circulation model (CGCM)]. The vector $\boldsymbol{\beta}$ contains the scaling factors that adjust the amplitude of each those signal patterns (also called fingerprints) to best match the observed amplitude, and \mathbf{u} is a realization of internal climate variability. Vector \mathbf{u} is assumed to be a realization of a Gaussian random vector (see below for discussion). Long “control” simulations with CGCMs, that is, simulations without anomalous external forcing, are typically used to estimate the internal climate variability and the resulting uncertainty in scaling factors $\boldsymbol{\beta}$.

Inferences about detection and attribution in the standard approach are then based on hypothesis testing. For detection, this involves testing the null hypothesis that the amplitude of a given signal is consistent with zero (if this is not the case, it is detected); attribution is assessed using the *attribution consistency test* (Allen and Tett 1999; see also Hasselmann 1997), which evaluates the null hypothesis that the amplitude $\boldsymbol{\beta}$ is a vector of units (i.e., the model signal does not need to be rescaled to match the observations). A complete attribution assessment accounts for competing mechanisms of climate change as completely as possible, as discussed by Mitchell et al. (2001). Increasingly, Bayesian approaches are used as an alternative to the standard approach. In Bayesian approaches, inferences are based on a posterior distribution that blends evidence from the observations with independent prior information that is represented by a prior distribution [e.g., Berliner et al. 2000; Schnur and Hasselmann 2004; Lee et al. 2005; see International Ad Hoc Detection and Attribution Group (2005) for a more complete discussion and results]. Since Bayesian approaches can incor-

porate multiple lines of evidence and account elegantly for uncertainties in various components of the detection and attribution effort, we expect that they will be very helpful for variables with considerable observational and model uncertainty.

As we move toward detection and attribution studies on smaller spatial and temporal scales and with non-temperature variables, new challenges arise that are related to noise and uncertainty in signal patterns, dealing with non-Gaussian variables and facing data limitations. These are now discussed.

a. Data requirements for detection and attribution

The observations analyzed in a detection approach should cover a long enough time period to distinguish an emerging anthropogenic signal, typically at least 20 yr, or better, 50 yr. Longer records generally allow for a more powerful detection of the anthropogenic signal from the background of natural variability, but the time period is limited by available observed data and samples for climate variability. The observed record also needs to be as homogeneous in time as possible; that is, free from artifacts due to changes in temporal sampling, instrument bias, instrument exposure or location, observing procedures, and processing algorithms.

Time-dependent biases for long time period temporal sampling (e.g., monthly, seasonal, and annual) have been addressed more frequently and effectively than biases associated with short temporal sampling (hourly and daily). However, analysis of climate extremes requires high-resolution temporal sampling. Difficulties arise from diurnal biases of temperature that are difficult to completely eliminate (see, e.g., DeGaetano 1999; Vose et al. 2003) and from corrections for short-duration precipitation integrations (hourly or less) versus longer time integrations (daily and monthly; Groisman et al. 1999).

Data availability is still limited, particularly in very high latitudes and the Tropics (see http://www.ncdc.noaa.gov/img/climate/research/2005/feb/map_prdp_02_2005_pg.gif). Also, there remains considerable data that are inaccessible in many developing and some developed countries. A U.S. program to rescue long-term data that are not electronically accessible (the U.S. Climate Data Modernization Program) is now working with other countries and the World Meteorological Organization (WMO) to help fill these gaps. It has already lead to new data being made available worldwide. Supporting information about instrument status and the observing environment is critical to derive appropriate corrections for time-dependent biases. Therefore, it is very important to also maintain and rescue metadata.

Detection methods may be helpful in prioritizing where observational data would be most useful to constrain model climate change fingerprints (see, e.g., Groisman et al. 2005). More needs to be done in that regard.

Reanalysis data are dynamically complete and can provide a valuable source of data for studying climate variability. However, at present, inhomogeneities in time, particularly during the time of transition to the satellite era, make these products problematic to use for detection (e.g., Chelliah and Ropelewski 2000). Limiting the analysis to the better-constrained satellite era, and analyzing data from several reanalyses, particularly more recent ones, can circumvent some of these problems (see Santer et al. 2003b; Gillett et al. 2003b), although caution is still needed.

b. Addressing error and noise in model-simulated patterns

Because CGCMs simulate natural internal variability as well as the response to external forcing, the CGCM simulated climate signals need to be averaged across an ensemble of simulations. Even then, signal estimates will contain remnants of the climate's natural internal variability unless the ensemble size is very large. The presence of this noise in the signal may bias ordinary least squares estimates of β downward, particularly for signals that have small signal-to-noise ratios (such as signals from natural forcing or other anthropogenic forcings in the twentieth century). This can be addressed by estimating β with a total least squares algorithm (Allen and Stott 2003). Further processing of signals or fingerprints (see Santer et al. 1996; Hegerl et al. 1996) may be needed to reduce the amount of noise for variables and spatial scales that are more strongly affected by climate variability.

Model-simulated signals also invariably contain uncertainties associated with errors in models (such as imperfect treatment of clouds, e.g.) and forcings. Detection and attribution results are sensitive to this uncertainty as demonstrated when results from different models and different forcing assumptions are compared (e.g., Santer et al. 1996; Hegerl et al. 2000; Allen et al. 2006). A first estimate of the combined model error and forcing uncertainty can be based on combining data from simulations forced with different estimates of radiative forcings, and simulated with different models. Gillett et al. (2002) demonstrate that such multimodel fingerprints lead to a more convincing attribution of observed warming between greenhouse gas and sulfate aerosol forcing. Taylor (K. Taylor 2005, personal communication) shows that averages from multiple models often outperform individual models in simulations of mean climate and variability.

For a complete understanding of the effects of forcing and model uncertainty, and a full representation of both uncertainties in detection and attribution approaches (as suggested by Hasselmann 1997), both forcing and model uncertainties need to be explored fully and separately. Using very large ensembles of models with perturbed parameters will improve the model error estimate (see Allen and Stainforth 2002; Murphy et al. 2004). However, if models share common errors, the estimate of model uncertainty will be biased low. It is therefore important to maintain true diversity in climate models used worldwide.

Also, appreciation of the complexities of the numerous types of anthropogenic and natural forcings is growing rapidly. Additional climate forcings have been identified recently, such as several types of aerosols, changes in land use, urbanization, and irrigation practices (e.g., Dolman et al. 2003; Bonan 1999; Charney 1975; Hahmann and Dickinson 1997). The importance of these forcings will vary between climate variables and spatial scales. For example, while land use change is thought to have a relatively small effect on globally or hemispherically averaged temperature (e.g., Matthews et al. 2004), it can have substantial effects locally (e.g., Baidya and Avissar 2002) and may therefore be important for the detection of regional climate change.

While forcing uncertainty affects the results of estimating contributions of external forcing to observed changes, detection methods can also provide help to constrain the magnitude of external forcings if their space–time signature is known (“top-down” forcing estimates; see, e.g., Anderson et al. 2003).

c. *Estimates of internal climate variability*

One of the primary concerns with current optimal fingerprinting techniques is related to the dependence upon models for estimates of internal variability. There are at least two prospects for improving our confidence in these estimates.

First, the paleoclimate community continues to make impressive progress in the reconstruction and interpretation of the climate record of the last 1–2 millennia (e.g., Jones and Mann 2004), although uncertainties remain (von Storch et al. 2004). However, the variability in paleoreconstructions is a convolution of internal climate variability, additional noise from proxy data, sampling uncertainty due to incomplete coverage of paleodata, and the climate response to uncertain external forcing. A comparison of this variability with unforced internal climate variability in climate models is not straightforward. One step toward such a comparison is comparing the residual variability in paleoclimatic reconstructions after removing effects from external forc-

ing (e.g., Hegerl et al. 2006) with variability in control simulations; or, alternatively, comparing the variability in simulations of the last millennium with proxy data (e.g., Tett et al. 2006). Studies of the last millennium also help to better understand climate response to natural forcing.

Second, the CGCMs that are used for climate change research are also increasingly being used for seasonal and longer-range prediction—that is, for use in initial value problems rather than external forcing response problems. Prediction skill at seasonal to interannual time scales is not necessarily an indicator of a model’s potential skill in simulating the response to external forcing. However, prediction research provides an understanding of the circumstances under which we can make skillful seasonal to interannual forecasts, and it can help to validate the mechanisms that provide that skill, thus increasing confidence in estimates of internal variability from CGCMs. This should also provide further insights into the large-scale feedback mechanisms that determine the climate’s sensitivity to forcing, and the nature of its transient response to that forcing (e.g., Boer et al. 2004, 2005; Boer 2004), since these mechanisms are also likely an important source of predictive skill on seasonal to interannual time scales.

d. *Linearity*

Another, but substantially smaller, concern is the “linear” model that is used predominantly in climate change detection research.¹ This model assumes that the responses to the various external agents (anthropogenic and natural) that are thought to have influenced the climate of the past century add linearly. There is little evidence to suggest that the response has not been additive on global scales. However, additivity may not continue to hold well on smaller space or time scales or in the future, and biogeochemical feedback mechanisms may cause nonadditive feedbacks on radiative forcing (e.g., Cox et al. 2000). A breakdown of additivity would pose a problem for the use of detection methods to constrain model projections of future climate, although it is possible to address this in the context of existing methods.

e. *Non-Gaussian variables and extremes*

A third area of concern is the extension of detection techniques so that they can be used to evaluate changes

¹ The word *linear* is used in a statistical sense in this context—it indicates linear scaling of the model-simulated space–time climate change signal. This use of the word *linear* does not describe the nature of the climate change signals that enter into the analysis—those signals may well evolve in a nonlinear fashion in time.

in quantities that are not inherently Gaussian, such as the detection and attribution of change in the frequency and intensity of extreme events.² This will be a challenge because signal-to-noise ratios are expected to be low. There are two fundamental challenges.

The first challenge is methodological and not inherently difficult. Inferences in current optimal fingerprinting methods can be understood as based on a “likelihood function.” The form of that function, and thus the method of inference, is derived from the “link” between the climate change signals and the observations, $\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{u}$, and the assumption that the errors are Gaussian. Research is already underway where the relationship between the observations and the signal is more complex than the simple equation above, and where the distribution function is replaced with one that is more appropriate for extremes (e.g., Kharin and Zwiers 2005; Wang et al. 2004; Zhang et al. 2005).

However, there are some additional and more difficult challenges in the detection of externally forced change in extremes. These include continuing challenges in resolving the scaling issues that hinder the comparison of CGCM simulated extremes with observed extremes (which will be discussed in section 5), and a lack of consensus between models on the simulation of present-day extremes (Kharin et al. 2005). Nonetheless, there is a pressing need for information in this area, and thus the detection community will increasingly venture into this area of research.

3. Large-scale change at the surface and in the atmosphere

a. Attribution of twentieth-century warming to causes

The conclusion of the third Intergovernmental Panel on Climate Change (IPCC) assessment report from detection and attribution studies was that “most of the observed warming over the last 50 years is likely to have been due to the increase in greenhouse gas concentrations” (Mitchell et al. 2001). This conclusion has been largely based on results using multiple regressions of observed surface air temperature onto fingerprints of greenhouse gas, sulfate aerosol or combined anthropo-

genic nongreenhouse gas emissions, and natural forcing (solar and/or volcanic forcing separately, or both combined). The effect of various uncertainties in detection and attribution results, such as forcing or model uncertainty as discussed above, is summarized in the term “likely.” Detection and attribution results from global surface temperature data will need to be updated with improved model versions, better estimates of forcing, and more complete estimates of uncertainty in order to better quantify and narrow the remaining uncertainty in detection results. A further issue that is being addressed but needs more work is observational uncertainty during the first half of the twentieth century (Smith and Reynolds 2003).

Progress has been made in understanding differences between surface and tropospheric temperature trends. The climate response to anthropogenic forcing in the vertical profile of temperature trends is characterized by stratospheric cooling and tropospheric warming. Such a climate response has been detected in radiosonde data since the 1960s (e.g., Santer et al. 1996; Tett et al. 1996; Allen and Tett 1999), even if only lower-tropospheric temperatures are considered (Thorne et al. 2003). Cooling of the stratosphere and warming of the troposphere leads to an increase in tropopause height, where clear anthropogenic and natural signals can be detected in a range of reanalysis data (Santer et al. 2003b).

The apparent lack of significant warming in the lower troposphere over the satellite era has raised concerns over the validity of estimates of surface warming (Christy et al. 2001; Christy and Norris 2004) or the ability of climate models to simulate the vertical coherence in temperature (e.g., Hegerl and Wallace 2002). This problem is discussed in International Ad Hoc Detection and Attribution Group (2005) and has been the subject of a U.S. Climate Change Science Program Synthesis Report (Karl et al. 2006). For understanding trends in satellite measurements of the upper troposphere, the influence of the stratosphere on that measurement needs to be considered. Recent analyses suggest that trends in surface and tropospheric temperature are consistent with how we expect them to vary according to the physics of the atmosphere if this stratospheric influence (and its temperature trends associated with stratospheric ozone depletion) as well as observational uncertainty in satellite data are considered (see Mears et al. 2003; Fu et al. 2004). The trends are also no longer inconsistent with model-simulated trends if observational uncertainty and natural forcing is considered (Santer et al. 2003a). However, the uncertainty in satellite data processing needed to be fully understood in order to yield an improved best guess and uncer-

² There are fewer distributional concerns with most current applications of the optimal fingerprinting approach, regardless of whether the variable of interest is temperature, precipitation, or some other quantity. This is because almost all studies have applied the technique to data that are composed of space–time averages computed over long periods of time (e.g., a decade) and large regions (e.g., $10^\circ \times 10^\circ$ or larger latitude–longitude boxes). According to the central limit theorem, these quantities should have distributions that are close to Gaussian.

tainty range for satellite-derived tropospheric temperature trends (Karl et al. 2006). This example demonstrates a need for improved operation of satellite and in situ observing systems for monitoring climate.

b. Changes in global circulation and precipitation

The atmospheric circulation is driven by differential heating across the globe, and as external forcing perturbs these heating rates, it is natural to expect the atmospheric circulation to change in response (see, e.g., Palmer 1999). However, there is no widely accepted theory to describe how it is likely to change. As discussed in International Ad Hoc Detection and Attribution Group (2005), positive trends in the Northern and Southern Annular Modes have recently been observed (Hurrell 1996; Thompson et al. 2000; Thompson and Solomon 2002; Gillett et al. 2003a). The surface circulation is well-characterized by sea level pressure, which has the advantage that it is well observed and exhibits a high degree of spatial homogeneity. Gillett et al. (2003b) used detection and attribution methods to compare simulated and observed trends in sea level pressure and found a detectable response to a combined greenhouse gas and sulfate aerosol forcing using three different observational datasets and the mean simulated response from four models [the Second Hadley Centre Coupled Ocean–Atmosphere GCM (HadCM2), Third HadCM (HadCM3), CGCM1, and CGCM2; note that of these models only HadCM3 has no flux corrections]. These results have now been extended to include the 40-Yr European Centre for Medium-Range Weather Forecasts (ECMWF) Re-Analysis (ERA-40). Although the pattern of simulated and observed change was similar, Gillett et al. (2003b) found that the magnitude of the observed sea level pressure change is substantially larger than that simulated in several climate models. Figure 1a shows changes in winter sea level pressure over the period 1958–98 from the ERA-40 dataset compared to the mean response simulated by four climate models (Fig. 1b). The simulated pattern of sea level pressure trends is similar to that from the reanalysis, but the magnitude is much smaller. This result is confirmed by Fig. 1c, which shows regression coefficients of sea level pressure change from several observed and reanalysis datasets against a multimodel mean simulated response to greenhouse gas and sulfate aerosol increases. The scaling factor (see section 2) is always significantly greater than one, indicating that the observed response is larger than that simulated by the models.

Why do climate models fail to predict the correct magnitude of sea level pressure changes? One reason may be that the studies discussed above do not include

all the relevant external climate forcings. Stott et al. (2001) examine integrations of HadCM3 forced with all of the principal external forcings—greenhouse gas changes, sulfate aerosol changes, solar irradiance changes, volcanic aerosol, and stratospheric ozone depletion—and find that they do not simulate the recently observed North Atlantic Oscillation (NAO) increase. However, Gillett and Thompson (2003) examined the response to stratospheric ozone depletion in a model with high vertical resolution and found that realistic December–February trends in geopotential height over the Southern Hemisphere were simulated in response to ozone depletion. They also noted that simulation of these trends required high vertical resolution, explaining why they were not simulated by Stott et al. (2001). These results thus suggest that part of the discrepancy between simulated and observed circulation changes in the Southern Hemisphere noted by Gillett et al. (2003b) may be due to ozone depletion. However, the discrepancy over the Northern Hemisphere cannot be explained in this way.

Shindell et al. (1999) argue that the tropospheric circulation response to greenhouse gas increases is remotely forced from the stratosphere and that a high model upper boundary is necessary in order to simulate a realistic sea level pressure response to greenhouse gas increases, but their findings were not reproduced in a model with higher horizontal resolution (Gillett et al. 2002). Other authors have suggested that the North Atlantic Oscillation response to greenhouse gas increases is indirectly forced by changes in sea surface temperatures (Rodwell et al. 1999; Hoerling et al. 2001), but while some studies with prescribed sea surface temperatures are able to simulate changes in the NAO that are correlated with those that have been observed, none has yet been able to simulate the magnitude of the observed trend. Thus the reason for the difference in amplitude of the observed and simulated sea level pressure trends remains unknown.

How might we reconcile this difference between observed and simulated sea level pressure changes? First, it is important to identify and characterize sources of uncertainty in the observational datasets. Sufficiently long instrumental records of sea level pressure only exist for limited areas of the globe, thus we must either restrict our analysis to only these well-observed regions or use sea level pressure derived from reanalyses. The National Centers for Environmental Prediction (NCEP) reanalysis exhibits larger negative trends in sea level pressure in the poorly observed Antarctic, which are not fully reproduced in the recent ERA-40 reanalysis, suggesting that the NCEP reanalysis trends may be overestimates there. A detection analysis applied to sea

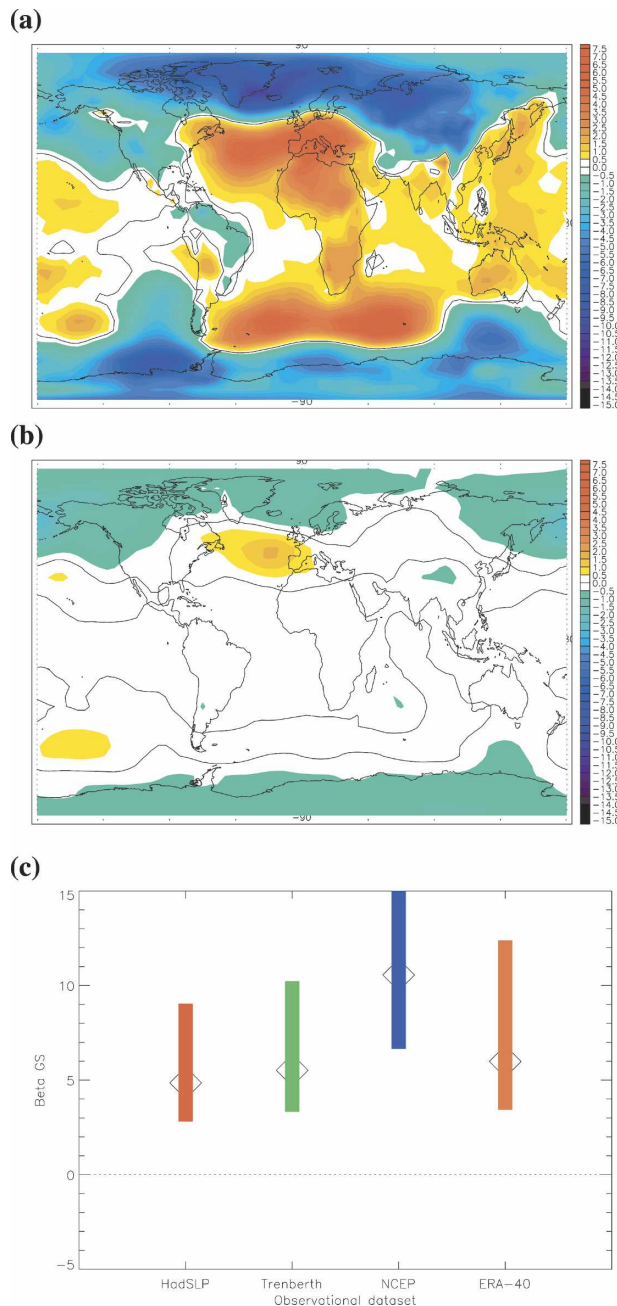


FIG. 1. (a) “Observed” and (b) simulated sea level pressure trends (same color scale, in hPa). The December–February sea level pressure trends over the period 1958–98 are shown (a) for ERA-40 and (b) for the mean of the simulated response to greenhouse gas and sulfate aerosol changes from four climate models (CGCM1, CGCM2, HadCM2, and HadCM3; adapted from Gillett et al. 2003b). (c) Regression coefficients (scaling factor β ; see section 2) of different observed sea level pressure changes against changes simulated in response to greenhouse gas and sulfate aerosol increases. The first three bars show results derived using a four-model mean of simulated 1948–98 sea level pressure changes, and three observational datasets: the Hadley Centre sea level pressure (HadSLP) is a gridded dataset of measurements, defined only where observations are present; Trenberth is a synthesis of analyses, only defined northward of 20°N; and NCEP is a global reanalysis. The fourth bar shows results using ERA-40 reanalysis over the period 1958–98 (adapted from Gillett et al. 2003b).

level pressure over the North Atlantic region (20°–80°N and 0°–60°W) over the period 1908–98 using the Trenberth analyses (Trenberth and Paolino 1980) indicated no better agreement between models and analysis-based sea level pressure data than the over the 1958–98 period, but further analysis of historical data may help to better constrain uncertainties.

We also need to examine climate models to understand why they are in disagreement with observations, if the observed trends prove correct. For example, it is likely that the sea level pressure response is sensitive to the parameterizations used in a model. By making use of “perturbed physics” ensembles (e.g., Allen 2003b), in which physical parameterizations are systematically perturbed in a large ensemble of integrations, it may be possible to identify the model parameters to which circulation changes are most sensitive and that lead to a more realistic simulation of historical sea level pressure changes. This area of disagreement between models and observations may therefore ultimately prove useful in constraining model physics.

As with atmospheric circulation changes, we also expect the hydrological cycle to respond to changes in external forcing of the climate system. Mitchell et al. (1987) argue that precipitation changes are controlled primarily by the energy budget of the troposphere: the latent heat of condensation being balanced by radiative cooling. Externally forced warming of the troposphere enhances the local cooling rate, thereby increasing precipitation, but this may be partly offset by a decrease in the efficiency of the cooling due to greenhouse gas increases (Allen and Ingram 2002; Yang et al. 2003; Lambert et al. 2004). Allen and Ingram (2002) demonstrate that the ensemble mean land average precipitation simulated by HadCM3 is significantly correlated with observed land average precipitation over the 1945–98 period, essentially detecting the influence of natural external forcing on precipitation. A similar result was obtained using all-forcings simulations of the Parallel Climate Model (PCM; Fig. 2; Gillett et al. 2004b). Consistent with this, Lambert et al. (2004) demonstrate that the response to shortwave forcing is detectable in observations, whereas the response to longwave forcing is not. These results therefore suggest that natural forcings such as volcanic aerosol and solar irradiance changes are likely to have had a larger influence on mean changes of total precipitation during the twentieth century than greenhouse gas changes, which is consistent with simulations of the response to volcanic forcing (Robock and Liu 1994). Gillett et al. (2004b) demonstrate that there is a detectable volcanic influence in terrestrial precipitation over the past 50 yr, using simu-

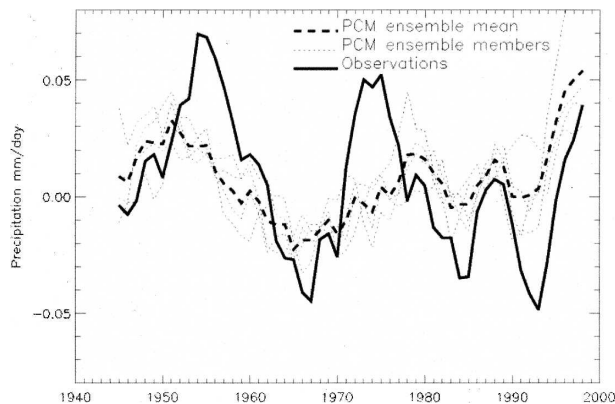


FIG. 2. Simulated and observed anomalies in terrestrial precipitation over the period 1945–98. The model data are from a four-member ensemble of simulations of the NCAR PCM integrated with estimates of anthropogenic, solar, and volcanic forcing. A 5-yr running mean was applied to suppress short-term natural variability, such as ENSO (taken from Gillett et al. 2004b).

lations of the PCM, although the model appears to underestimate the volcanic response.

Owing to the limited sensitivity of precipitation to greenhouse gas changes and the relatively small change in forcing over the observed period, Allen and Ingram (2002) argue that hydrological sensitivity (the change in mean total precipitation in response to a doubling of CO_2) is not well constrained by available observations. A perfect model study that examined the detectability of precipitation changes in simulations of the PCM with natural and anthropogenic forcing also suggested that the response to greenhouse gas forcing should not yet be detectable in total mean precipitation (Ziegler et al. 2003) and that model uncertainty should make detection of annual precipitation changes difficult (Hegerl et al. 2004 note that changes in some aspects of extreme precipitation may be detectable earlier; see below). However, detection and attribution techniques are likely to be useful in examining the hydrological response to natural forcings, particularly volcanoes. In these cases, we may be able to use these techniques to answer the question of whether observed and simulated precipitation responses are consistent, and in the context of a perturbed physics ensemble, these techniques may be used to constrain model parameters by comparison with observations (to the extent that observations provide a constraint given their uncertainties). This in turn may help to constrain our predictions of future precipitation changes.

A major impediment in detection of anthropogenic influences on precipitation is that global estimates of precipitation are not available, particularly before the satellite era. Decadal changes recorded by satellite

measurements of rainfall are still uncertain. Station-based datasets over land are incomplete, even during the past 50 yr, and are also affected by observational uncertainties (Houghton et al. 2001).

4. Changes in the ocean

There is an increasing amount of observational evidence for changes within the ocean, both at regional and global scales (e.g., Bindoff and Church 1992; Wong et al. 1999; Wong et al. 2001; Dickson et al. 2001; Curry et al. 2003; Levitus et al. 2001; Aoki et al. 2005). Many of the observed changes in the ocean are from studies of the heat storage (Ishii et al. 2003; White et al. 2003; Willis et al. 2004; Levitus et al. 2005). These studies all show that the global heat content of the oceans has been increasing since the 1950s. For the period 1993–2003, this increase is between 0.7 and 0.86 W m^{-2} . The longer-term average increase in heat content (1955–98) over the 0–3000-m layer of the ocean is 0.2 W m^{-2} or 0.037°C . These observed changes in ocean heat content are consistent with model-simulated changes in state-of-the-art coupled climate models, which can be detected and attributed to anthropogenic forcing (e.g., Barnett et al. 2001; Levitus et al. 2001; Reichert et al. 2002). However, total ocean heat content is affected by observational sampling uncertainty (Gregory et al. 2004). Since the ocean is a major source of uncertainty in future climate change (see Houghton et al. 2001), attempting to detect and quantify ocean climate change in variables focusing on ocean physics, such as water mass characteristics, will increase confidence in large-scale simulations of climate change in the ocean and our ability to simulate future ocean changes.

The water mass characteristics of the relatively shallow Sub-Antarctic Mode Water (SAMW) and the subtropical gyres in the Indian and Pacific basins since the 1960s have been changing. In most studies differences between earlier historical data (mainly from the 1960s) with more recent World Ocean Circulation Experiment (WOCE) data in the late 1980s and 1990s show that the SAMW is cooler and fresher on density surfaces (Bindoff and Church 1992; Johnson and Orsi 1997; Bindoff and McDougall 1994; Wong et al. 2001; Bindoff and McDougall 2000), indicative of a subduction of warmer waters [see Bindoff and McDougall (1994) for an explanation of this counterintuitive result]. These water mass results are supported by the strong increase in heat content in the Southern Hemisphere midlatitudes across both the Indian and Pacific Oceans during the 1993–2003 period (Willis et al. 2004). While most studies of the SAMW water mass properties have shown a cooling and freshening on density surfaces in the Indian

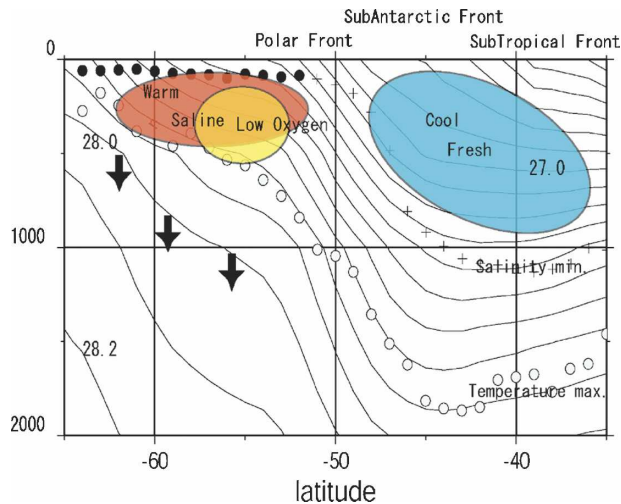


FIG. 3. Schematic summary of the meridional water mass differences between data from the 1990s and the 1950s to the 1970s in the Southern Ocean from South Africa to south of Australia (45°E–90°W). [Printed with permission of *Geophys. Res. Lett.* from Aoki et al. (2005).]

and Pacific Oceans, the most recent repeat of the WOCE Indian Ocean section along 32°S in 2001 found a warming and salinity increase on density surfaces (indicative of subduction of cooler waters) in the shallow thermocline (Bryden et al. 2003). This result emphasizes the need to understand the processes involved in decadal oscillations in the subtropical gyres. Note, however, that the denser waters masses below 300 m showed the same trend in water mass properties that had been reported earlier (Bindoff and McDougall 2000). Further evidence of the large-scale freshening and cooling of SAMW (Fig. 3) comes from an analysis of six meridional WOCE sections and three Japanese Antarctic Research Expedition sections from South Africa to 150°E. These sections were compared with historical data extending from the Subtropical Front (~35°S) to the Antarctic Divergence (~60°S), and from South Africa eastward to the Drake Passage. In almost all sections a cooling and freshening of SAMW has occurred consistent with the subduction of warmer surface waters observed over the same period, summarized in Fig. 3.

The salinity minimum water in the North Pacific has freshened and in the southern parts of the Atlantic, Indian, and Pacific Oceans there has also been a corresponding freshening of the salinity minimum layer. The Atlantic freshening at depth is also supported by direct observations of a freshening of the surface waters (Curry et al. 2003). Taken together these changes in the Atlantic and North Pacific suggest a global increase in the hydrological cycle (and flux of freshwater into the

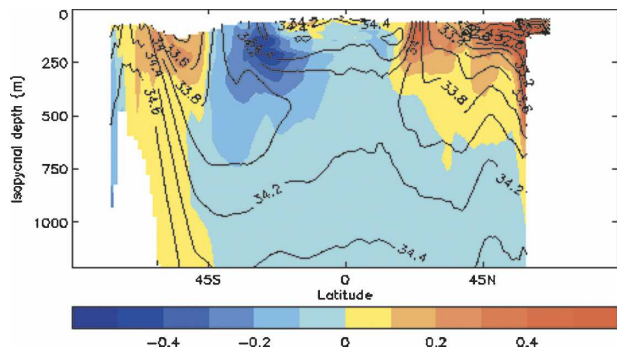


FIG. 4. Indo-Pacific zonal average potential temperature change (°C) on isopycnals from HadCM3 anthropogenic climate change experiment between 1959–68 and 1998–98. The changes on the isopycnals have been regridded onto pressure surfaces using the mean depth of each isopycnal surface. Contours show the zonal average salinity field (psu). [Printed with permission of *J. Climate* from Banks and Bindoff (2003).]

oceans including melt waters from ice caps and sea ice) at high latitudes in the source regions of these two water masses (Wong et al. 1999). To the south of the Subantarctic Front, there is a very coherent pattern of warming and salinity increase on density surfaces <500 m (Fig. 3). This pattern of warming and salinity increase on isopycnals from 45°E to 90°W is consistent with the warming and/or freshening of surface waters (see Bindoff and McDougall 1994). Figure 3 summarizes the observed differences in the Southern Ocean, showing the cooling and freshening on density surfaces of SAMW north of the Subantarctic Front, freshening of Antarctic Intermediate Water, and warming and salinity increase of the Upper-Circumpolar Deep Water south of the Subantarctic Front.

These observed changes are broadly consistent with simulations of warming and changes in precipitation minus evaporation. Banks and Bindoff (2003) identified a zonal mode (or fingerprint) of difference in water mass properties in the anthropogenically forced simulation of the HadCM3 model between the 1960s and 1990s (Fig. 4). This fingerprint identified in HadCM3 is strikingly similar to the observed differences in water mass characteristics (Fig. 3) in the Southern Hemisphere. In the HadCM3 climate change simulation, the strength of the zonal mode in the Indo-Pacific Ocean tends to become stronger and increasingly significant from the 1960s onward. Its strength exceeds the 5% significance level 40% of the time, while this happens only occasionally (<5% of the time) in the 600-yr control simulation. This result suggests that the zonal signature of climate change for the Indo-Pacific basin (and Southern Ocean) is distinct from the modes of variability and suggests that the anthropogenic change can be

separated from internal ocean variability. The similarity of observed and simulated water mass changes suggests that such changes can already be observed. For a quantitative detection approach, the relatively sparse sampling of ocean data needs to be emulated in models.

The Southern Ocean is an important source of world's global water masses and thermohaline circulation. Banks and Wood (2002) concluded from their analysis of the HadCM3 model results that the geographic regions with the greatest signal-to-noise ratio for detecting climate changes trends were from water masses that mainly originate from or in the Southern Ocean with short residence times such as SAMW. By contrast, the North Atlantic was considered less suitable for climate change detection because of its greater internal variability in this model.

5. Detecting anthropogenic changes in impact-relevant variables

a. Toward detecting regional changes

Regional and local changes in climate have a large impact on society. Recently, it has been shown that an anthropogenic climate change signal is detectable in continental-scale regions using surface temperature changes over the twentieth century (Karoly et al. 2003; Stott 2003; Zwiers and Zhang 2003; Karoly and Braganza 2005). It has also been shown that most of the observed warming over the last 50 yr in six continental-scale regions (including North America, Eurasia, and Australia) was likely to be due to the increase in greenhouse gases in the atmosphere (Stott 2003). However, it becomes harder to detect climate change at decreasing spatial scales, and scaling factors may become more model dependent. This tendency is illustrated in Fig. 5, which is based on the approach by Zwiers and Zhang (2003). The authors use the Canadian climate model to show that greenhouse gas and sulfate aerosol climate change can be detected in the observed warming in North America and Eurasia over the twentieth century. As the spatial scales considered become smaller, it can be seen that the uncertainty in estimated signal amplitudes (as demonstrated by the size of the vertical bars) becomes larger, reducing the signal-to-noise ratio in detection and attribution results (see also Stott and Tett 1998). Since the signal-to-noise ratio depends on the local level of natural variability and the size of the anthropogenic signal, results vary between regions, such as between Eurasia and North America. The figure also illustrates that most of the results hold if the variance of internal climate variability in the control simulations is doubled [by enhancing anomalies of the control simulation by a factor of $\sqrt{2}$; see Fig. 5b]. This increases

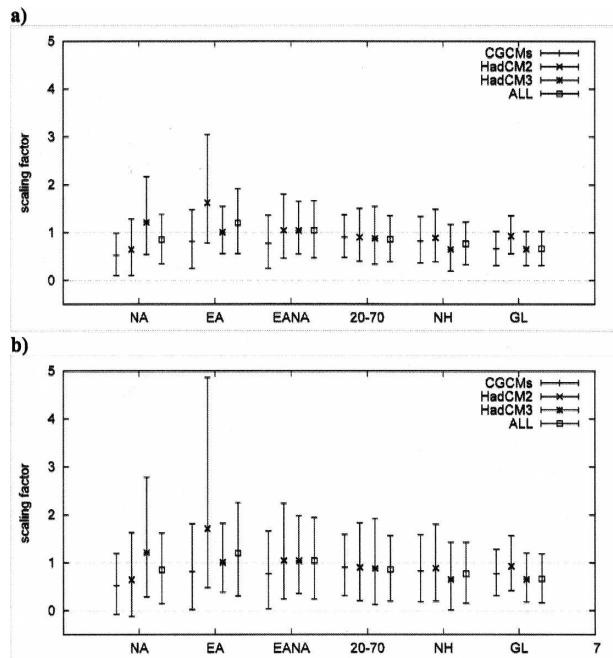


FIG. 5. (a) Detection of greenhouse gas and aerosol climate change in observed data from seven nested regions (after Zwiers and Zhang 2003): NA, EA, EANA, 30–70, NH, and GL represent the North American continent, the Eurasian continent, the North American and Eurasian continents combined, Northern Hemisphere midlatitudes (30°–70°N), Northern Hemisphere, and the globe, respectively. Simulations from CGCM1, CGCM2, HadCM2, and HadCM3 have been used as fingerprints. ALL corresponds to the results where data from all four models have been combined in the fingerprint and estimate of climate variability. The vertical axis shows the scaling factors by which model data need to be scaled to match observations, and their 90% confidence intervals. A scaling factor that is significantly larger than unity indicates detection of climate change in observations, overlap of the confidence interval with unity indicates that simulated and observed changes are consistent. (b) Same as in (a), except that the variance in the control simulations has been doubled. Note that a significant detection of the greenhouse gas and aerosol signal still holds everywhere if multimodel detection (ALL) is used.

our confidence in the detection result, since estimates of internal climate variability based on models are still uncertain (see section 2).

A different approach to detection of regional temperature change uses indices of area-average minimum and maximum surface temperature variations in the North American region (Karoly et al. 2003) and in the Australian region (Karoly and Braganza 2005) calculated from observations and a number of different climate models. Results show that recent climate change in those regions could not be explained by natural variability alone and was consistent with the response to anthropogenic forcing (Fig. 6).

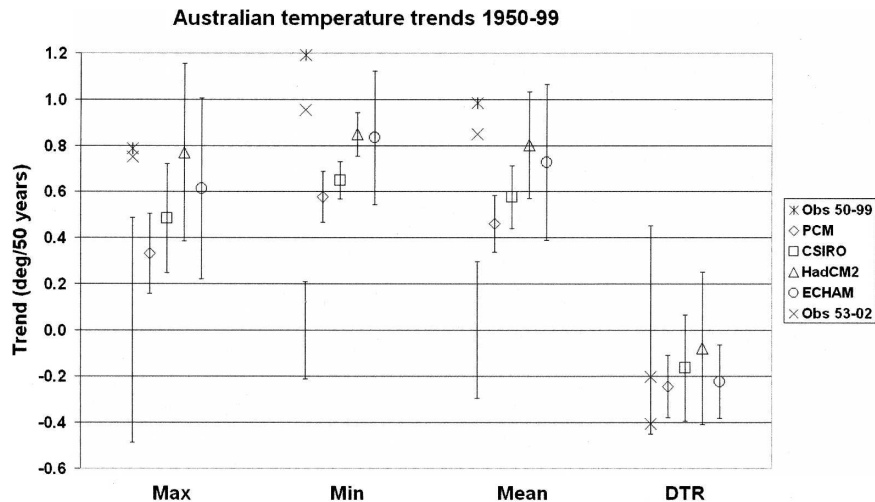


FIG. 6. Trends in Australian area-averaged min, max, and mean temperatures, and daily temperature range (DTR) from anthropogenically forced model simulations and observations over 1950–99. The error bars on the model trends are the 90% confidence intervals for the ensemble-mean trends (based on control variability, taking ensemble size into account, omitted for models with short control runs available to authors). The error bars about zero at the location of the observed trends are the uncertainties in the trend estimates due to natural internal climate variability, as simulated by the models (based on CSIRO and PCM models, which were the only models with long control runs with DTR data available). The trends in mean temperature are inconsistent with naturally forced runs (from Karoly and Braganza 2005; Braganza et al. 2004).

The successful attribution of continental-scale climate change to anthropogenic forcing, as demonstrated in the results discussed above, can also be used to provide probabilistic estimates of future climate change at regional scales (in a similar manner as done for global scales; see, e.g., Stott and Kettleborough 2002).

Detection of regional climate change is very relevant for attributing impacts of climate change to external forcing. Gillett et al. (2004a) demonstrate a detectable anthropogenic influence on Canadian fire season temperature. They go on to detect the influence of anthropogenic climate change on forest area burnt, using a simple statistical model. This result links observed impacts directly to external forcing. Such an approach will become increasingly important for understanding climate change impacts, such as changes in ecosystems.

However, the prospects of successful attribution of observed temperature change at local scales (such as at a single station) are limited in the near future, as the magnitude of local temperature, and even more so, rainfall, variability is generally much larger than any regional greenhouse climate change signal. The spatial scale at which a detectable anthropogenic signal can be identified is likely to decrease over time, as the magnitude of the projected greenhouse climate signal increases.

b. Extreme events

Perhaps one of the most unexpected developments in the area of climate change detection and attribution is the recent focus on extreme climate events. Certainly, from the perspective of climate impacts extreme weather and climate events are very important, but until recently it was not expected that they would exhibit detectable anthropogenic signals beyond a shift due to changes in climate means in the near future. However, the central Europe heat wave during the summer of 2003, which is estimated to be a very extreme event in the context of long station records, is consistent with hypothesized increases in temperature variability and hence greater likelihood of extremes (Schär et al. 2004).

Results from climate model simulations suggest that the tails of the distribution of daily temperature data will change differently from seasonal mean data, suggesting that a separate detection of changes in temperature extremes is worthwhile. Figure 7 shows that two climate models simulate a stronger change in European cold winter days than in winter means, narrowing the future temperature distribution in a manner consistent with simulated changes in circulation, while the distribution of daily maximum temperature widens, leading to stronger hot extremes (Hegerl et al. 2004).

Climatological data show that the most intense pre-

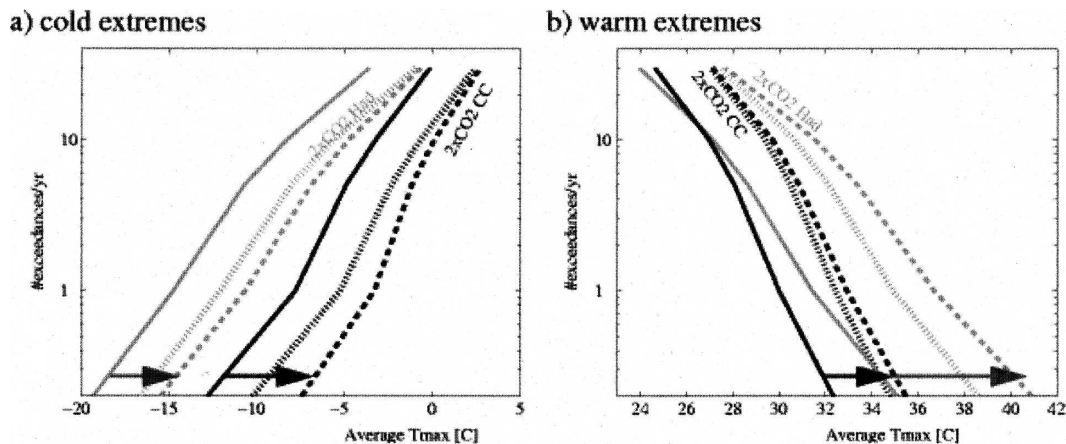


FIG. 7. Change in moderate to rare daily maximum temperature in central Europe for (a) cold extremes and (b) warm extremes. On the horizontal axis, temperature is given ($^{\circ}\text{C}$); the vertical axis indicates (on a log scale) the number of most extreme days per year that have been averaged (ranging from 30 to 0.2), with the average of the 30 hottest and coldest days of a year given on top and the average of the hottest/coldest days in a 5-yr period at the bottom of the diagram. The solid line indicates the mean value for the model's present-day climate (1975–95 in a climate change simulation), dashed for future climate (2040–60). Changes are given for two different climate models, CGCM2 (gray) and HadCM3 (black). Note that the more rare extremes change more strongly for both cold and warm extremes in both models (indicated by the arrows); the dotted line indicates the respective seasonal mean change for comparison. [From Hegerl et al. (2004).]

cipitation occurs in warm regions (Fig. 8a). Also, higher temperatures lead to an increase in the water holding capacity of the atmosphere, and hence to a greater proportion of total precipitation in heavy and very heavy events (Karl and Trenberth 2003). Therefore, all climate models analyzed to date show on average an increase in extreme precipitation events as global temperatures increase (Houghton et al. 2001; Semenov and Bengtsson 2002; Allen and Ingram 2002; Hegerl et al. 2003), with global increases in extreme precipitation exceeding increases in mean precipitation. Groisman et al. (1999) has demonstrated empirically, and Katz (1999) theoretically, that as precipitation increases a greater proportion falls in heavy and very heavy events if the frequency remains constant. Figure 8b illustrates that observed decadal trends in rainfall tend to show stronger changes in extreme than mean rainfall. Although measurement uncertainties in these regional changes are considerable, the probability of 16 out of 16 regions showing stronger absolute changes in extremes than means by chance is very small. Note that this result, which applies to the 90th percentile of daily precipitation, is not inconsistent with model results that suggest that the magnitude of very rare events, such as the 20-yr extreme event, will increase almost everywhere with increasing temperature.

These findings draw attention to the necessity of closer examination of the changes in precipitation extremes and attempts to detect changes and attribute them to anthropogenic forcing. However, there are a

number of difficulties to address before such a detection and attribution attempt becomes feasible.

First, as mentioned in the methods section, a comparison between observed and simulated changes in climate extremes requires a comparison of data that represent different spatial scales: while the typical global climate model grid box is on the order of one or several hundreds of kilometers wide, the observations represent point observations by individual stations. Therefore, a direct quantitative comparison between observed and simulated extremes is not feasible, and it is important to develop area-averaged changes in extreme precipitation (Groisman et al. 2005). A large number of stations are needed to provide reliable estimates of area-averaged precipitation (e.g., McCollum and Krajewski 1998; Osborn and Hulme 1997). Data from reanalysis projects (e.g., ERA-40 reanalysis, Simmons et al. 2005; or the updated NCEP reanalysis, Kanamitsu et al. 2002) may be useful since they are more readily comparable to model data, but rainfall in these products is not well constrained by observations [see Kharin and Zwiers (2000) for extreme and Widmann et al. (2003) for mean rainfall]. On the other hand, if reanalysis rainfall extremes are driven by parameterizations, we might be able to learn from the success or failure of different reanalysis products about model parameterizations that improve the simulation of extreme rainfall.

Today, station-based observations are the most reliable data for detection and attribution of climate change in rainfall extremes, but as the time series and

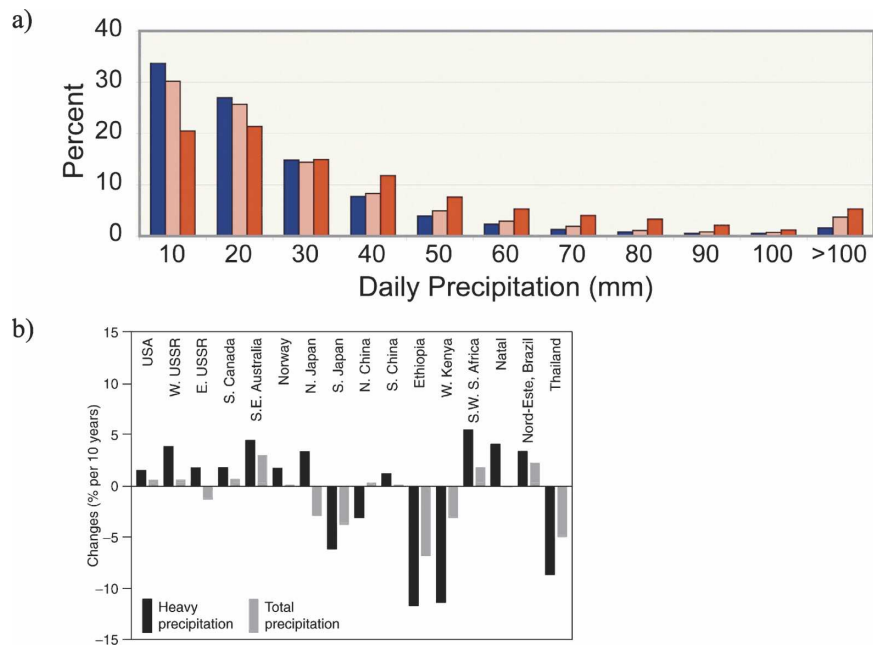


FIG. 8. (a) Climatology of the intensity of daily precipitation in 10 mm day^{-1} categories for different temperature regimes: blue bar: -3° to 19°C ; pink bar: 19° to 29°C ; dark red bar: 29° to 35°C , based on 51, 37, and 12 worldwide stations, respectively. By selection, all stations have the same seasonal mean precipitation amount of $230 \pm 5 \text{ mm}$. [From Karl and Trenberth (2003).] (b) Linear trends ($\% \text{ decade}^{-1}$) of heavy precipitation (above 90th percentile) and total precipitation during the rainy season over various regions of the globe. Seasons for each region usually span at least 50 yr. Note that the magnitude of the changes in heavy precipitation frequencies are always higher than changes in mean precipitation frequencies. [From Easterling et al. (2000).]

accuracy of remote sensing data increases, the blend of these different types of data will become increasingly important for comparison with climate model simulations. Station data still require additional work for daily and possibly hourly resolution data, for integration into global datasets and assessment for time-dependent biases caused by systematic changes in observing procedure or instruments. Fortunately, the impacts of such systematic changes in precipitation observations appear to be strongest for light precipitation measurements and affect less the measurement of heavy and very heavy precipitation (Groisman et al. 1999). Other inhomogeneities, such as changes in station location, may still affect heavy rainfall, though these are less spatially coherent.

A second difficulty in the detection of changes in extremes is that the term “climate extreme” encompasses a range of events that typically cause impacts. These range from frequent events such as midlatitude frost days to extremely rare and devastating events. Consequently, a large range of indices documenting extreme events has been proposed and applied (see Meehl et al. 2000; Frich et al. 2002). This different use

of indices for extremes has so far made a comparison between results of model and observational studies of extremes difficult (see Houghton et al. 2001). Examples for indices of extremes include the most extreme event over a period of time, such as a year. This index may be interesting by itself (Hegerl et al. 2004) or can be used to fit an extreme value distribution that allows us to estimate extreme events with long return characteristics (see Zwiers and Kharin 1998; Kharin and Zwiers 2000, 2005; Wehner 2004). Other indices of extremes are defined as exceedances of a threshold for extreme events, such as the 90th percentile of climatological temperature or rainfall. Exceedances of thresholds benefit from extensive statistical literature on their properties. However, their application to climate variables with strong seasonal cycle, such as temperature, leads to unanticipated problems. Thresholds that are based on estimated percentiles of climatological temperature are affected by sampling error. This error leads to systematic differences in exceedance rates between the climatological base period and the period outside, causing substantial biases in trends in extremes (Zhang et al. 2005). These can be circumvented if extremes indices are pro-

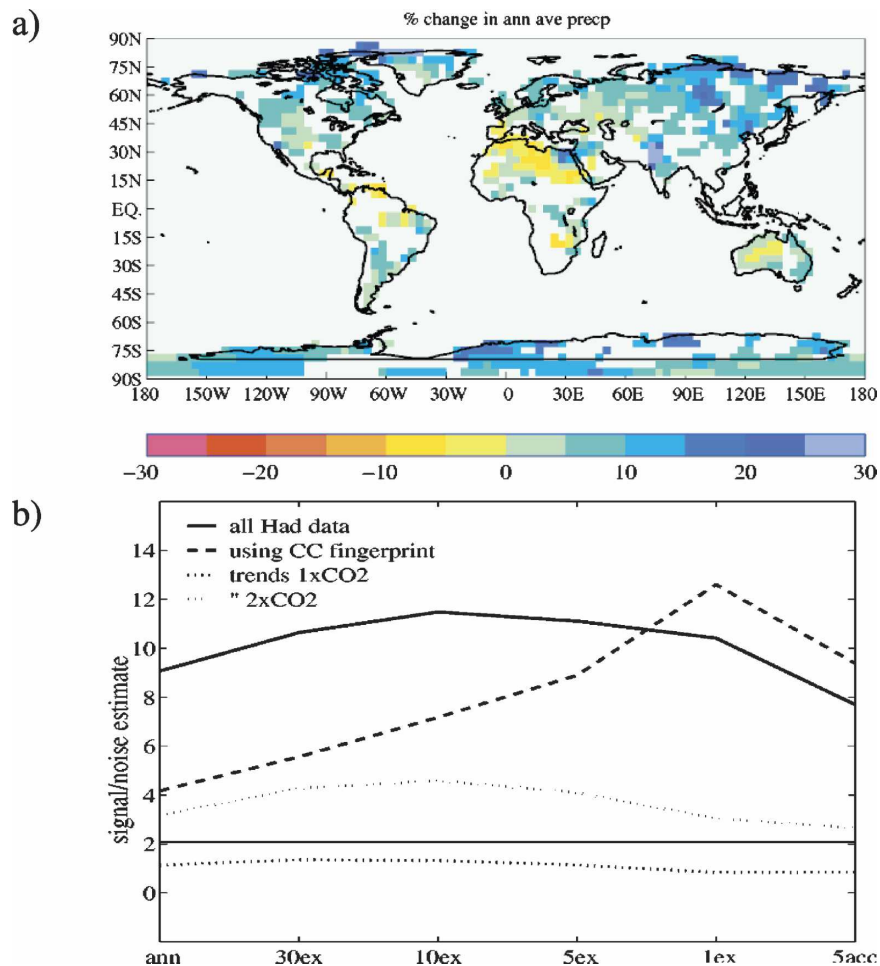


FIG. 9. (a) Common model fingerprint for the percent change in the wettest day of the year between the present time and the time of CO₂ doubling from two climate models (changes are averaged between the two climate models CGCM2 and HadCM3 and shown only where changes are consistent; in percent). (b) Signal-to-noise ratio for changes in extreme precipitation estimated from model data only. The horizontal axis indicates results from annual total ("ann"), the average of 30, 10, 5, and 1 wettest days ("30ex" to "1ex") and the strongest 5-day accumulation ("5acc") if detected using data from HadCM3 only ("all Had data"; solid) and using the fingerprint from CGCM2 ("using CC fingerprint"; dashed; results are similar if models are exchanged). The latter results show that changes in heavy precipitation should be more detectable between both models. [Further details can be found in Hegerl et al. (2004).]

cessed differently. This example demonstrates that indices for climate extremes must be very carefully evaluated for their statistical properties, their applicability to climatologically different regions, and their robustness. Data from climate models are very valuable to test the properties of indices, since they are abundant and relatively homogeneous.

A further consideration in the choice of indices is that indices for more rare and extreme events will be more poorly sampled than indices of events that occur more frequently. This decreased sampling will almost certainly lead to a decrease in signal-to-noise ratio for detection. However, for extremes that occur at least

once a year, this decrease in signal-to-noise ratio appears quite small for temperature or precipitation compared to seasonal mean data. If uncertainty in the spatial fingerprint of climate change in models is considered, changes in annual rainfall extremes may actually be more robustly detectable than changes in annual total rainfall (Hegerl et al. 2004; Fig. 9). This is caused by the above-mentioned stronger increases (in percent of climatological values) for extreme than annual total rainfall, which leads to a more robustly detectable pattern of general increase of extreme rainfall. In contrast, annual total rainfall shows a model-dependent pattern of increases and decreases.

This should encourage attempts to detect changes in extremes. A first attempt was based on the Frich et al. (2002) indices, using fingerprints from atmospheric model simulations with fixed sea surface temperature and a bootstrap method for significance testing (Kiktev et al. 2003). Their results indicate that patterns of simulated and observed rainfall extremes bear little similarity for the indices they selected, in contrast to the similarity of trends depicted by Groisman et al. (2005). In contrast, some observed changes in temperature extremes can be detected and attributed to greenhouse gas forcing (Christidis et al. 2005).

c. *Attributing individual extreme events probabilistically*

A new challenge for the detection and attribution community is quantifying the impact of external climate forcing on the probability of specific weather events. Detection and attribution studies to date have tended to focus on properties of the climate system that can be considered as deterministic. For example, the studies reviewed by the Houghton et al. (2001) attributing large-scale temperature changes were all based on the underlying statistical model of a deterministic change with superimposed climate noise. The combination of observational uncertainty and natural internal variability means that we cannot be completely sure what the externally driven 100-yr change in global temperatures has been, but can estimate a best guess and uncertainty range for the underlying anthropogenic temperature change from observed trends.

This distinction between the observed change in actual temperatures and the underlying change in expected temperatures is largely of academic interest when addressing global temperature trends, because the level of internal variability in 50- or 100-yr temperature trends is lower than the externally driven changes. This distinction becomes much more important when we consider changes in precipitation or extreme weather events. Nevertheless, even for these noisier variables, studies have tended to consider underlying deterministic changes in diagnostics such as expected occurrence frequency as the legitimate subject of attribution statements, rather than addressing the actual extreme events themselves. Indeed, in popular discussions of the climate change issue, it is frequently asserted that it is impossible in principle to attribute a single event in a chaotic system to external forcing.

Allen (2003a), Stone and Allen (2005), and Stott et al. (2004) argue that quantitative attribution statements can be made regarding individual events if they are couched in terms of the contribution of external forcing to the risk (i.e., the probability) of an event of (or

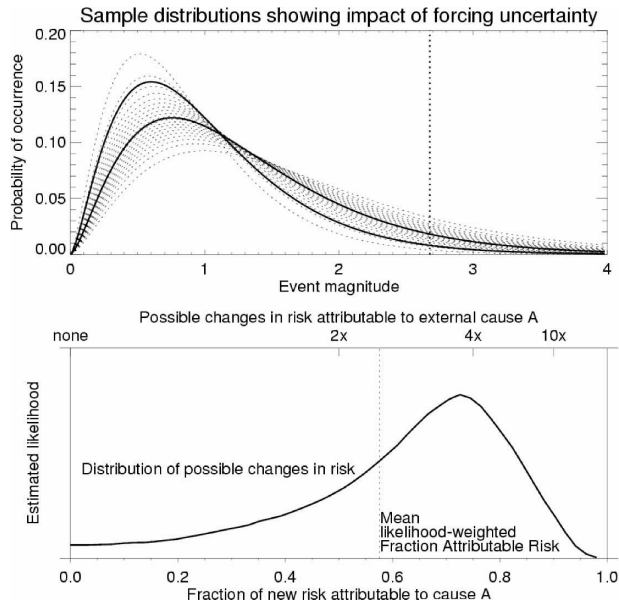


FIG. 10. (top) Probability distribution function (PDF) of a hypothetical weather event as a function of magnitude. In this example, the narrower distribution corresponds to the case without cause A (e.g., if humans had not influenced the climate system), whereas the wider distribution corresponds to the real-world case with cause A. The likelihood of an extreme event greater than the dotted line is doubled when cause A is present. The dotted lines show the range of possible PDFs in the real world including uncertainty in climate change. The uncertainty is large enough here that the probability of an extreme event could have decreased with cause A. Generally there will also be uncertainty in the reference PDF, which is not shown here. (bottom) PDF of the fraction of the risk of the extreme event attributable to cause A (see Allen 2003a; Stone and Allen 2005). Because of our uncertainty in the distribution in (a), the PDF here is wide enough to include a small probability that the risk has in fact decreased (i.e., less than 0 on the bottom axis).

greater than) the observed magnitude. This point is illustrated conceptually in Fig. 10. Figure 10a shows how the distribution of a hypothetical climate variable (precipitation at a given location, e.g.) might alter under climate change, with a narrower distribution changing to a broader distribution, increasing the risk of an event exceeding a given threshold. For assessing changes in risk, it will be necessary to account for uncertainty in how the distribution has changed: in this case, there is a 5% chance that the risk of exceeding the threshold has actually declined. Figure 10b (from Allen 2003a) shows how results from such probabilistic analyses can be summarized, showing a histogram of changes in risk resulting from the imposed external forcing (top axis) and the fraction attributable risk (FAR) due to that forcing (bottom axis). The FAR is an established concept in epidemiological studies for attribution of cause and effect in stochastic systems. It has been applied to attributing a part of the probability of a heat wave as

observed in central Europe in 2003 to anthropogenic forcing (Stott et al. 2004).

6. Recommendations and conclusions

Results of detection and attribution studies in surface and atmospheric temperature and ocean heat content show consistently that a large part of the twentieth-century warming can be attributed to greenhouse gas forcing. We need to continue to attempt estimating the climate response to anthropogenic forcing in different components of the climate system, including the oceans, atmosphere, and cryosphere. We also need to more fully assess all the components of the climate system for their sensitivity to climate change signals and their signal-to-noise ratios for climate change, and synthesize estimates of anthropogenic signals from different climate variables. Also, detection studies are now starting to focus on spatial scales and variables that are important for climate change impacts. All these efforts raise both familiar and new questions for climate research.

For example, the detection and attribution of climate change requires long observed time series free from nonclimate-related time-dependent biases. For the analysis of extreme events it is also important that quality control routines do not weed out true extreme events. Blended remote sensing and in situ data, if quality controlled also with regard to extremes, may become very useful to overcome spatial sampling inadequacies. Since every source of data is subject to observational uncertainty, climate records that are based on different observing systems and analysis methods are important for quantifying and decreasing the uncertainty in detection and attribution results. Lessons learned from microwave satellite data, global land surface temperatures, and sea surface temperatures show that our initial estimates of uncertainty from a single dataset are often too low. Therefore, a high priority must be placed on adequate estimation of error, including time-dependent biases.

For reducing uncertainties in detection and attribution results we also need to keep improving our understanding and estimates of historical anthropogenic and natural radiative forcings, particularly those with largest uncertainties such as black carbon, effect of aerosols on clouds, or solar forcing. As the spatial scale upon which detection and attribution efforts focus decreases, forcings that are of minor importance globally, such as land use change, may become more important and need to be considered.

Furthermore, our understanding of model uncertainty needs to be improved, and more complete esti-

mates of model error need to be included in detection and attribution approaches. Both ensembles of models with perturbed parameters (e.g., Allen and Stainforth 2002; Murphy et al. 2004) and true diversity in CGCMs used worldwide are important to sample model uncertainty. Aspects of climate change where there is a significant discrepancy between model simulation and observation, such as the magnitude of changes in annular modes or in the fingerprint of anthropogenic sea level pressure change, need to be understood.

Furthermore, different components of the climate system present their own challenges. In the oceans, it is important to exploit the signatures of climate change in both water mass properties, heat and freshwater, sea level, and in other ocean tracers, such as oxygen concentration, together to more reliably detect and attribute climate change and evaluate ocean model performance. The advantage of exploring water mass variations on density surfaces in addition to inventories of heat and freshwater storage, is that water mass changes largely reflect changes in the surface forcing and are less prone to noise introduced by mesoscale eddies. Furthermore, water mass changes on density surfaces do not contribute to sea level rise and thus provide information about changes within the water column that are independent from sea level measurements.

In the atmosphere, detectable global precipitation changes in response to volcanism may be useful to evaluate simulated changes in the hydrological cycle even before greenhouse gas-induced changes in precipitation become detectable. Also, changes in extreme precipitation may become detectable before changes in total precipitation. Furthermore, the probability of an individual extreme event with and without greenhouse warming can be estimated to assess how much global warming contributes to changes in the risk of a particular extreme event.

We conclude that while the anthropogenic signal continues to emerge from the background of natural variability in more components of the climate system, and on decreasing spatial scales, detection and attribution efforts will be vital to provide a rigorous comparison between model-simulated and observed change in both the atmosphere and oceans. Where climate change is detected and attributed to external forcing, detection results can be used to constrain uncertainties in future predictions based on observed climate change. Where attribution fails due to discrepancies between simulated and observed change, this provides an important encouragement to revisit climate model and observational uncertainty.

Acknowledgments. GCH was supported by NSF Grants ATM-0002206 and ATM-0296007, by NOAA Grant NA16GP2683 and NOAA's Office of Global programs, DOE in conjunction with the Climate Change Data and Detection element, and by Duke University. ERA-40 data used in this study have been obtained from the ECMWF data server. We thank Jesse Kenyon and Daithi Stone for help and discussion, and two anonymous reviewers and Tom Smith for their helpful comments.

REFERENCES

- Allen, M. R., 2003a: Liability for climate change. *Nature*, **421**, 891–892.
- , 2003b: Possible or probable? *Nature*, **425**, 242.
- , and S. F. B. Tett, 1999: Checking for model inconsistency in optimal fingerprinting. *Climate Dyn.*, **15**, 419–434.
- , and W. J. Ingram, 2002: Constraints on future changes in climate and the hydrologic cycle. *Nature*, **419**, 224–232.
- , and D. A. Stainforth, 2002: Towards objective probabilistic climate forecasting. *Nature*, **419**, 228.
- , and P. A. Stott, 2003: Estimating signal amplitudes in optimal fingerprinting, Part I: Theory. *Climate Dyn.*, **21**, 477–491.
- , —, J. F. B. Mitchell, R. Schnur, and T. Delworth, 2000: Quantifying the uncertainty in forecasts of anthropogenic climate change. *Nature*, **407**, 617–620.
- , and Coauthors, 2006: Quantifying anthropogenic influence on recent near-surface temperature. *Surv. Geophys.*, in press.
- Anderson, T. L., R. J. Charlson, S. E. Schwartz, R. Knutti, O. Boucher, H. Rodhe, and J. Heintzenberg, 2003: Climate forcing by aerosols—A hazy picture. *Science*, **16**, 1103–1104.
- Aoki, S., N. L. Bindoff, and J. A. Church, 2005: Interdecadal watermass changes in the Southern Ocean from 30°E and 160°E. *Geophys. Res. Lett.*, **32**, L07607, doi:10.1029/2004GL022220.
- Baidya, R. S., and R. Avissar, 2002: Impact of land use/land cover change on regional hydrometeorology in the Amazon. *J. Geophys. Res.*, **107**, 8037, doi:10.1029/2000JD000266.
- Banks, H. T., and R. A. Wood, 2002: Where to look for anthropogenic climate change in the ocean. *J. Climate*, **15**, 879–891.
- , and N. L. Bindoff, 2003: Comparison of observed temperature and salinity changes in the Indo-Pacific with results from the coupled climate model HadCM3: Processes and mechanisms. *J. Climate*, **16**, 156–166.
- Barnett, T. P., D. W. Pierce, and R. Schnur, 2001: Detection of anthropogenic climate change in the world's oceans. *Science*, **292**, 270–274.
- Berliner, L. M., R. A. Levine, and D. J. Shea, 2000: Bayesian climate change assessment. *J. Climate*, **13**, 3805–3820.
- Bindoff, N. L., and J. A. Church, 1992: Warming of the water column in the southwest Pacific Ocean. *Nature*, **357**, 59–62.
- , and T. J. McDougall, 1994: Diagnosing climate change and ocean ventilation using hydrographic data. *J. Phys. Oceanogr.*, **24**, 1137–1152.
- , and —, 2000: Decadal changes along an Indian Ocean section at 32°S and their interpretation. *J. Phys. Oceanogr.*, **30**, 1207–1222.
- Boer, G. J., 2004: Long timescale potential predictability in an ensemble of coupled climate models. *Climate Dyn.*, **23**, 29–44.
- , B. Yu, S.-J. Kim, and G. Flato, 2004: Is there observational support for an El Niño-like pattern of future global warming? *Geophys. Res. Lett.*, **31**, L06201, doi:10.1029/2003GL018722.
- , K. Hamilton, and W. Zhu, 2005: Climate sensitivity and climate change under strong forcing. *Climate Dyn.*, **24**, doi:10.1007/s00382-004-0500-3.
- Bonan, G. B., 1999: Frost followed the plow: Impacts of deforestation on the climate of the United States. *Ecol. Appl.*, **9**, 1305–1315.
- Braganza, K., D. J. Karoly, A. C. Hirst, P. Stott, R. J. Stouffer, and S. F. B. Tett, 2004: Simple indices of global climate variability and change: Part II—Attribution of climate change during the 20th century. *Climate Dyn.*, **22**, 823–838.
- Bryden, H. L., E. L. McDonagh, and B. A. King, 2003: Changes in ocean water mass properties: Oscillation or trends? *Science*, **300**, 2086–2088.
- Charney, J. G., 1975: Dynamics of deserts and drought in the Sahel. *Quart. J. Roy. Meteor. Soc.*, **101**, 193–202.
- Chelliah, M., and C. F. Ropelewski, 2000: Reanalysis-based tropospheric temperature estimates: Uncertainties in the context of global climate change detection. *J. Climate*, **13**, 3187–3205.
- Christidis, N., P. A. Stott, S. Brown, G. C. Hegerl, and J. Caesar, 2005: Detection of changes in temperature extremes during the second half of the 20th century. *Geophys. Res. Lett.*, **32**, L20716, doi:10.1029/2005GL023885.
- Christy, J. R., and W. B. Norris, 2004: What may we conclude about tropospheric temperature trends? *Geophys. Res. Lett.*, **31**, L06211, doi:10.1029/2003GL019361.
- , D. E. Parker, S. J. Brown, I. Macadam, M. Stendel, and W. B. Norris, 2001: Differential trends in tropical sea surface and atmospheric temperatures. *Geophys. Res. Lett.*, **28**, 183–186.
- Cox, P. M., R. A. Betts, C. D. Jones, S. A. Spall, and I. J. Tatterdell, 2000: Acceleration of global warming due to carbon-cycle feedbacks in a coupled climate model. *Nature*, **408**, 184–187.
- Curry, R., B. Dickson, and I. Yashayaev, 2003: Ocean evidence of a change in the freshwater balance of the Atlantic over the past four decades. *Nature*, **426**, 826–829.
- DeGaetano, A. T., 1999: A method to infer observation time based on day-to-day temperature variations. *J. Climate*, **12**, 3443–3456.
- Dickson, B., J. Hurrell, N. L. Bindoff, A. P. S. Wong, B. Arbic, B. Owens, S. Imawaki, and I. Yashayaev, 2001: The world during WOCE. *WOCE Conference Volume*, G. Siedler, J. A. Church, and J. Gould, Eds., Academic Press, 557–583.
- Dolman, A. J., A. Verhagen, and C. A. Rovers, Eds., 2003: *Global Environmental Change and Land Use*. Kluwer Academic, 210 pp.
- Easterling, D. R., G. A. Meehl, C. Parmesan, S. Changnon, T. R. Karl, and L. O. Mearns, 2000: Climate extremes: Observations, modeling and impacts. *Science*, **289**, 2068–2074.
- Forest, C. E., P. H. Stone, A. P. Sokolov, M. R. Allen, and M. D. Webster, 2002: Quantifying uncertainties in climate system properties with the use of recent observations. *Science*, **295**, 113–117.
- Frich, P., L. V. Alexander, P. Della-Marta, B. Gleason, M. Haylock, A. M. G. Klein-Tank, and T. Peterson, 2002: Observed coherent changes in climatic extremes during the second half of the twentieth century. *Climate Res.*, **19**, 193–212.
- Fu, Q., C. M. Johanson, S. G. Warren, and D. J. Seidel, 2004: Contribution of stratospheric cooling to satellite-inferred tropospheric temperature trends. *Nature*, **249**, 55–58.

- Gillett, N. P., and D. W. J. Thompson, 2003: Simulation of recent Southern Hemisphere climate change. *Science*, **302**, 273–275.
- , F. W. Zwiers, A. J. Weaver, G. C. Hegerl, M. R. Allen, and P. A. Stott, 2002: Detecting anthropogenic influence with a multi-model ensemble. *Geophys. Res. Lett.*, **29**, 1970, doi:10.1029/2002GL015836.
- , H. Graf, and T. Osborn, 2003a: Climate change and the NAO. *The North Atlantic Oscillation*, *Geophys. Monogr.*, Vol. 134, Amer. Geophys. Union, 193–210.
- , F. W. Zwiers, A. J. Weaver, and P. A. Stott, 2003b: Detection of human influence on sea level pressure. *Nature*, **422**, 292–294.
- , A. J. Weaver, F. W. Zwiers, and M. D. Flannigan, 2004a: Detecting the effect of human-induced climate change on Canadian forest fires. *Geophys. Res. Lett.*, **31**, L18211, doi:10.1029/2004GL020876.
- , —, F. W. Zwiers, and M. F. Wehner, 2004b: Detection of volcanic influence on global precipitation. *Geophys. Res. Lett.*, **31**, L12217, doi:10.1029/2004GL020044.
- Gregory, J. M., H. T. Banks, P. A. Stott, J. A. Lowe, and M. D. Palmer, 2004: Simulated and observed decadal variability in ocean heat content. *Geophys. Res. Lett.*, **31**, L15312, doi:10.1029/2004GL020258.
- Groisman, P. Ya., and Coauthors, 1999: Changes in the probability of heavy precipitation: Important indicators of climatic change. *Climate Change*, **42**, 243–283.
- , R. W. Knight, D. R. Easterling, T. R. Karl, G. C. Hegerl, and V. N. Razuvayev, 2005: Trends in intense precipitation in the climate record. *J. Climate*, **18**, 1343–1367.
- Hahmann, A. N., and R. E. Dickinson, 1997: RCCM2-BATS model over tropical South America: Applications to tropical deforestation. *J. Climate*, **10**, 1944–1964.
- Hasselmann, K., 1979: On the signal-to-noise problem in atmospheric response studies. *Meteorology of Tropical Oceans*, D. B. Shaw, Ed., Royal Meteorological Society, 251–259.
- , 1997: Multi-pattern fingerprint method for detection and attribution of climate change. *Climate Dyn.*, **13**, 601–612.
- Hegerl, G. C., and J. M. Wallace, 2002: Influence of patterns of climate variability on the difference between satellite and surface temperature trends. *J. Climate*, **15**, 2412–2428.
- , H. von Storch, K. Hasselmann, B. D. Santer, U. Cubasch, and P. D. Jones, 1996: Detecting greenhouse-gas-induced climate change with an optimal fingerprint method. *J. Climate*, **9**, 2281–2306.
- , K. Hasselmann, U. Cubasch, J. F. B. Mitchell, E. Roeckner, R. Voss, and J. Waszkewitz, 1997: Multi-fingerprint detection and attribution of greenhouse-gas and aerosol-forced climate change. *Climate Dyn.*, **13**, 613–634.
- , P. Stott, M. Allen, J. F. B. Mitchell, S. F. B. Tett, and U. Cubasch, 2000: Detection and attribution of climate change: Sensitivity of results to climate model differences. *Climate Dyn.*, **16**, 737–754.
- , F. W. Zwiers, V. V. Kharin, and P. A. Stott, 2004: Detectability of anthropogenic changes in temperature and precipitation extremes. *J. Climate*, **17**, 3683–3700.
- , T. Crowley, M. Allen, W. T. Hyde, H. Pollack, J. Smerdon, and E. Zorita, 2006: Detection of human influence on a new, validated 1500-year temperature reconstruction. *J. Climate*, in press.
- Hoerling, M. P., J. W. Hurrell, and T. Y. Xu, 2001: Tropical origins for recent North Atlantic climate change. *Science*, **292**, 90–92.
- Houghton, J. T., Y. Ding, D. J. Griggs, M. Noguer, P. J. van der Linden, X. Dai, K. Maskell, and C. A. Johnson, Eds., 2001: *Climate Change 2001: The Scientific Basis*. Cambridge University Press, 881 pp.
- Hurrell, J. W., 1996: Influence of variations in extratropical wintertime teleconnections on Northern Hemisphere temperature. *Geophys. Res. Lett.*, **23**, 655–668.
- International Ad Hoc Detection and Attribution Group, 2005: Detecting and attributing external influences on the climate system: A review of recent advances. *J. Climate*, **18**, 1291–1314.
- Ishii, M., M. Kimoto, and M. Kachi, 2003: Historical subsurface temperature analysis with error estimates. *Mon. Wea. Rev.*, **131**, 51–73.
- Johnson, G. C., and A. H. Orsi, 1997: Southwest Pacific Ocean water-mass changes between 1968/69 and 1990/91. *J. Climate*, **10**, 306–316.
- Jones, P. D., and M. E. Mann, 2004: Climate over past millennia. *Rev. Geophys.*, **42**, RG2002, doi:10.1029/2003RG000143.
- Kanamitsu, M., W. Ebisuzaki, J. Woollen, S.-K. Yang, J. J. Hnilo, M. Fiorino, and G. L. Potter, 2002: NCEP–DOE AMIP-II Reanalysis (R-2). *Bull. Amer. Meteor. Soc.*, **83**, 1631–1643.
- Karl, T. R., and K. E. Trenberth, 2003: Modern global climate change. *Science*, **302**, 1719–1723.
- , S. J. Hassol, C. D. Miller, and W. L. Murray, Eds., 2006: Temperature trends in the lower atmosphere: Steps for understanding and reconciling differences. Climate Change Science Program and the Subcommittee on Global Change Research Report, Washington, DC.
- Karoly, D. J., and K. Braganza, 2005: Attribution of recent temperature changes in the Australian region. *J. Climate*, **18**, 457–464.
- , —, P. A. Stott, J. M. Arblaster, G. A. Meehl, A. J. Broccoli, and D. W. Dixon, 2003: Detection of a human influence on North American climate. *Science*, **302**, 1200–1203.
- Katz, R. W., 1999: Extreme value theory for precipitation: Sensitivity analysis for climate change. *Adv. Water Resour.*, **23**, 133–139.
- Kharin, V. V., and F. W. Zwiers, 2000: Changes in the extremes in an ensemble of transient climate simulations with a coupled atmosphere–ocean GCM. *J. Climate*, **13**, 3760–3788.
- , and —, 2005: Estimating extremes in transient climate change simulations. *J. Climate*, **18**, 1156–1173.
- , —, and X. Zhang, 2005: Intercomparison of near-surface temperature and precipitation extremes in AMIP-2 simulations, reanalyses, and observations. *J. Climate*, **18**, 5201–5223.
- Kiktev, D., D. Sexton, L. Alexander, and C. Folland, 2003: Comparison of modeled and observed trends in indices of daily climate extremes. *J. Climate*, **16**, 3560–3571.
- Lambert, F. H., P. A. Stott, M. R. Allen, and M. A. Palmer, 2004: Detection and attribution of changes in 20th century land precipitation. *Geophys. Res. Lett.*, **31**, L10203, doi:10.1029/2004GL019545.
- Lee, T. C. K., F. W. Zwiers, X. Zhang, G. C. Hegerl, and M. Tsao, 2005: A Bayesian climate change detection and attribution assessment. *J. Climate*, **18**, 2429–2440.
- Levitus, S., J. Antonov, J. Wang, T. L. Delworth, K. W. Dixon, and A. J. Broccoli, 2001: Anthropogenic warming of the Earth's climate system. *Science*, **292**, 267–270.
- , —, and T. Boyer, 2005: Warming of the world ocean, 1955–2003. *Geophys. Res. Lett.*, **32**, L02604, doi:10.1029/2004GL021592.
- Matthews, H. D., A. J. Weaver, K. J. Meissner, N. P. Gillett, and M. Eby, 2004: Natural and anthropogenic climate change:

- Incorporating historical land cover change, vegetation dynamics and the global carbon cycle. *Climate Dyn.*, **22**, 461–479.
- McCollum, J. R., and W. F. Krajewski, 1998: Uncertainty of monthly rainfall estimates from rain gauges in the Global Precipitation Climatology Project. *Water Resour. Res.*, **34**, 2647–2654.
- Mears, C. A., M. C. Schabel, and F. J. Wentz, 2003: A reanalysis of the MSU channel 2 tropospheric temperature record. *J. Climate*, **16**, 3650–3664.
- Meehl, G. A., F. Zwiers, J. Evans, T. Knutson, L. Mearns, and P. Whetton, 2000: Trends in extreme weather and climate events: Issues related to modeling extremes in projections of future climate change. *Bull. Amer. Meteor. Soc.*, **81**, 427–436.
- Mitchell, J. F. B., C. A. Wilson, and W. M. Cunningham, 1987: On CO₂ climate sensitivity and model dependence of results. *Quart. J. Roy. Meteor. Soc.*, **113**, 293–322.
- , D. J. Karoly, G. C. Hegerl, F. E. Zwiers, and J. Marengo, 2001: Detection of climate change and attribution of causes. *Climate Change 2001: The Scientific Basis*, J. T. Houghton et al., Eds., Cambridge University Press, 695–738.
- Murphy, J. M., D. M. H. Sexton, D. N. Barnett, G. S. Jones, M. J. Webb, M. Collins, and D. A. Stainforth, 2004: Quantification of modelling uncertainties in a large ensemble of climate change simulations. *Nature*, **430**, 768–772.
- Osborn, T. J., and M. Hulme, 1997: Development of a relationship between station and grid-box rainfall frequencies for climate model evaluation. *J. Climate*, **10**, 1885–1908.
- Palmer, T. N., 1999: A nonlinear dynamical perspective on climate prediction. *J. Climate*, **12**, 575–591.
- Reichert, K. B., R. Schnur, and L. Bengtsson, 2002: Global ocean warming tied to anthropogenic forcing. *Geophys. Res. Lett.*, **29**, 1525, doi:10.1029/2001GL013954.
- Robock, A., and Y. Liu, 1994: The volcanic signal in Goddard Institute for Space Studies three-dimensional model simulations. *J. Climate*, **7**, 44–55.
- Rodwell, M. J., D. P. Powell, and C. K. Folland, 1999: Oceanic forcing of the wintertime North Atlantic Oscillation and European climate. *Nature*, **398**, 320–323.
- Santer, B. D., and Coauthors, 1996: A search for human influences on the thermal structure in the atmosphere. *Nature*, **382**, 39–45.
- , and Coauthors, 2003a: Influence of satellite data uncertainties on the detection of externally-forced climate change. *Science*, **300**, 1280–1284.
- , and Coauthors, 2003b: Contributions of anthropogenic and natural forcing to recent tropopause height changes. *Science*, **301**, 479–483.
- Schär, C., P. L. Vidale, D. Lüthi, C. Frei, C. Häberli, M. A. Liniger, and C. Appenzeller, 2004: The role of increasing temperature variability in European summer heatwaves. *Nature*, **427**, doi:10.1038/nature02300.
- Schnur, R., and K. Hasselmann, 2004: Optimal filtering for Bayesian detection of climate change. *Climate Dyn.*, **24**, doi:10.1007/s00382-004-0456-3.
- Semenov, V. A., and L. Bengtsson, 2002: Secular trends in daily precipitation characteristics: Greenhouse gas simulation with a coupled AOGCM. *Climate Dyn.*, **19**, 123–140.
- Shindell, D. T., R. L. Miller, G. A. Schmidt, and L. Pandolfo, 1999: Simulation of recent northern winter climate trends by greenhouse-gas forcing. *Nature*, **399**, 452–455.
- Simmons, A. J., M. Hortal, G. Kelly, A. McNally, A. Untch, and S. Uppala, 2005: ECMWF analyses and forecasts of stratospheric winter polar vortex breakup: September 2002 in the Southern Hemisphere and related events. *J. Atmos. Sci.*, **62**, 668–689.
- Smith, T. M., and R. W. Reynolds, 2003: Extended reconstruction of global sea surface temperatures based on COADS data (1854–1997). *J. Climate*, **16**, 1495–1510.
- Stone, D. A., and M. R. Allen, 2005: The end-to-end attribution problem: From emissions to impacts. *Climatic Change*, **71**, 303–318.
- Stott, P. A., 2003: Attribution of regional-scale temperature changes to anthropogenic and natural causes. *Geophys. Res. Lett.*, **30**, 1728, doi:10.1029/2003GL017324.
- , and S. F. B. Tett, 1998: Scale-dependent detection of climate change. *J. Climate*, **11**, 3282–3294.
- , and J. A. Kettleborough, 2002: Origins and estimates of uncertainty in predictions of 21st century temperature rise. *Nature*, **416**, 723–726.
- , S. F. B. Tett, G. S. Jones, M. R. Allen, W. J. Ingram, and J. F. B. Mitchell, 2001: Attribution of twentieth century temperature change to natural and anthropogenic causes. *Climate Dyn.*, **17**, 1–21.
- , D. A. Stone, and M. R. Allen, 2004: Human contribution to the European heatwave of 2003. *Nature*, **432**, 610–614.
- Tett, S. F. B., J. F. B. Mitchell, D. E. Parker, and M. R. Allen, 1996: Human influence on the atmospheric vertical temperature structure: Detection and observations. *Science*, **274**, 1170–1173.
- , P. A. Stott, M. R. Allen, W. J. Ingram, and J. F. B. Mitchell, 1999: Causes of twentieth century temperature change near the earth's surface. *Nature*, **399**, 569–572.
- , and Coauthors, 2006: The impact of natural and anthropogenic forcings on climate and hydrology since 1550. *Climate Dyn.*, in press.
- Thompson, D. W. J., and S. Solomon, 2002: Interpretation of recent Southern Hemisphere climate change. *Science*, **296**, 895–899.
- , J. M. Wallace, and G. C. Hegerl, 2000: Annular modes in the extratropical circulation. Part II: Trends. *J. Climate*, **13**, 1018–1036.
- Thorne, P. W., and Coauthors, 2003: Probable causes of late twentieth century tropospheric temperature trends. *Climate Dyn.*, **21**, 573–591.
- Trenberth, K. E., and D. A. Paolino, 1980: The Northern Hemisphere sea-level pressure data set: Trends, errors and discontinuities. *Mon. Wea. Rev.*, **108**, 855–872.
- von Storch, H., E. Zorita, J. M. Jones, Y. Dimitriev, F. González-Rouco, and S. F. B. Tett, 2004: Reconstructing past climate from noisy data. *Science*, **306**, doi:10.1126/science.1096109.
- Vose, R. S., C. N. Williams, T. C. Peterson, T. R. Karl, and D. R. Easterling, 2003: An evaluation of the time of observation bias adjustment in the U.S. Historical Climatology Network. *Geophys. Res. Lett.*, **30**, 2046, doi:10.1029/2003GL018111.
- Wang, X. L., F. W. Zwiers, and V. R. Swail, 2004: North Atlantic ocean wave climate change scenarios for the twenty-first century. *J. Climate*, **17**, 2368–2383.
- Wehner, M. F., 2004: Predicted twenty-first-century changes in seasonal extreme precipitation events in the Parallel Climate Model. *J. Climate*, **17**, 4281–4290.
- White, W. B., M. D. Dettinger, and D. R. Cayan, 2003: Sources of global warming of the upper ocean on decadal period scales. *J. Geophys. Res.*, **108**, 3248, doi:10.1029/2002JC001396.
- Widmann, M., C. S. Bretherton, and E. P. Salathe, 2003: Statistical precipitation downscaling over the northwestern United

- States using numerically simulated precipitation as a predictor. *J. Climate*, **16**, 799–816.
- Willis, J. K., D. Roemmich, and B. Cornuelle, 2004: Interannual variability in upper-ocean heat content, temperature and thermosteric expansion on global scales. *J. Geophys. Res.*, **109**, C12036, doi:10.1029/2003JC002260.
- Wong, A., N. L. Bindoff, and J. A. Church, 1999: Large-scale freshening of intermediate waters in the Pacific and Indian Oceans. *Nature*, **400**, 440–443.
- , —, and —, 2001: Freshwater and heat changes in the North and South Pacific Oceans between the 1960s and 1985–94. *J. Climate*, **14**, 1613–1633.
- Yang, F., A. Kumar, M. E. Schlesinger, and W. Wang, 2003: Intensity of hydrological cycles in warmer climates. *J. Climate*, **16**, 2419–2423.
- Zhang, X., G. Hegerl, F. W. Zwiers, and J. Kenyon, 2005: Avoiding inhomogeneity in percentile-based indices of temperature extremes. *J. Climate*, **18**, 1641–1651.
- Ziegler, A. D., J. Sheffield, E. P. Maurer, B. Nijssen, E. F. Wood, and D. P. Lettenmaier, 2003: Detection of intensification in global- and continental-scale hydrological cycles: Temporal scale of evaluation. *J. Climate*, **16**, 535–547.
- Zwiers, F. W., and V. K. Kharin, 1998: Changes in the extremes of the climate simulated by CCC GCM2 under CO₂ doubling. *J. Climate*, **11**, 2200–2222.
- , and X. Zhang, 2003: Toward regional-scale climate change detection. *J. Climate*, **16**, 793–797.